# ENZYME POLYMORPHISM AND ADAPTATION IN ALPINE BUTTERFLIES[1]

GEORGE B. JOHNSON[2]

ABSTRACT

The high levels of enzyme polymorphism often detected by electrophoresis probably reflect on-going balancing selection. However, hypotheses of single gene heterosis (such as overdominance or environmental balance) do not seem adequate to account for this variation. To investigate this problem further, polymorphism was studied in Colorado butterflies of the genus *Colias*. *Colias* are easy to breed and study in the field and occupy a diverse array of habitats. Four species were studied: all were found to be highly polymorphic. Polymorphism at the α-glycerophosphate dehydrogenase locus occurred only when the species in question occupied a montane habitat; no polymorphism was seen in populations occupying alpine or lowland habitats. Three populations of *C. meadii* were examined in detail which encompass both alpine and montane habitats (crossing timberline); in one of them detailed demographic studies were carried out. In each such population, marked clines were observed in α-GPdH frequency, despite the swamping effects of migration within the population. At thirteen other loci a variety of different patterns of allele frequency are seen (some clinal, some uniform, some discontinuous), despite the fact that all loci were assayed in the same individuals. This result constitutes strong evidence of selection. There is clear evidence that particular alleles at the different loci preferentially associate together at specific locations along the cline, different assemblages occurring at different locations. A hypothesis is presented that these represent integrated metabolic phenotypes, and that the enzyme polymorphism is a multi-locus strategy to preserve that integration in a heterogeneous environment.

## INTRODUCTION

In the last decade it has become abundantly clear that levels of genetic variability detected by electrophoresis are very high in animal populations. However, the evolutionary significance of this variation is not clear. The difficulty is in understanding why there is so *much* of it. Most natural populations seem to be polymorphic at around a third of their enzyme loci, and over 10% of individuals are heterozygous at a typical enzyme locus (Johnson, 1973a; Selander & Johnson, 1973; Lewontin, 1974; Harris et al., 1974; Powell, 1975). This is far more genetic variability than theory had led us to expect. The disparity from expectation is not unlike Diogenes searching for one honest man—and finding hundreds! Among population geneticists there has been a lively discussion concerning the possibility that these very high levels of polymorphism are simply "noise," with no adaptively-important differences between alleles. I will not review that controversy here, except to say that I feel the weight of the evidence favors an adaptive interpretation. Rejecting the *null* hypothesis is only the first step, however. It remains to understand the biological meaning of this unexpectedly large amount of gene variation. It is with this problem in mind that I wish to describe some aspects of the population biology of the butterfly *Colias*. Detailed study of biochemical polymorphisms in this genus suggests approaches which may help to clarify the issue.

---

[1] CIW-DPB Publication Number 559.

[2] Department of Biology, Washington University, St. Louis, Missouri 63130. *Present address*: Carnegie Institution of Washington, Department of Plant Biology, 290 Panama Street, Stanford, California 94305.

ANN. MISSOURI BOT. GARD. 63: 248–261. 1976.

## Experimental Approach: A Single Locus Study

To assess the adaptive significance of genetic variation is not a trivial matter. It is not enough to simply observe patterns of allele frequency which correlate with some aspect of the natural environment, as many factors other than adaptation may generate such patterns: migration, founder effect, genetic drift in small populations, linkage to other loci under selection, all may have important effects. What is required is a well-defined empirical question designed to directly contrast adaptive values. To study polymorphic variation at a single gene locus, an ideal system would involve:

1. Polymorphism for an enzyme whose physiological function is well known.
2. An organism where genetic verification of allelism is possible.
3. A quantifiable environmental factor known to significantly influence the physiological function in question.
4. Demographically characterized populations.
5. Populations living in habitats which differ in terms of the chosen environmental factor.

In such a system it is possible to examine biochemical adaptation at a single locus directly (Clarke, 1975). One may ask whether there are indeed kinetic differences between alleles, whether the functioning of the allozymes is differentially affected by the environmental factor, and whether the polymorphic patterns are consistent with the habitat differences.

*Colias* provides such an experimental system. Butterflies may be raised in the laboratory (on hydroponically-germinated *Vicia* the life cycle of *C. eurytheme* is about a month) and pair-wise matings made to verify the mendelian segregation of variants. By marking the wings of live individuals one may carry out mark-release-recapture studies in a straight-forward manner, and thus learn the size and genetic structure of natural populations. In Colorado a variety of species of *Colias* occur, living in quite different habitats. As an organism for the approach outlined above, this butterfly thus seems a good choice.

To examine polymorphism at a single well-characterized locus I have chosen $\alpha$-glycerophosphate dehydrogenase ($\alpha$-GPdH). This enzyme performs in insects much the same function that LdH does in mammals: it acts to modulate the $NAD^+$/NADH redox level in the cell. During insect flight this is very important physiologically, as without a means of regenerating $NAD^+$ prolonged flight is impossible (thus *null* mutants at this locus in *Drosophila* are flightless).

The habitat-dependence of *Colias* flight has been examined in detail by Watt (1968). These butterflies act as thermodynamic "black boxes," flying only within a very narrow range of body temperature. This critical thermal range is typically above ambient air temperature in Colorado, and the butterflies rely on solar heating to raise their body temperatures to within the bounds of the thermal flight window. This is readily observable in the field, where, when a cloud covers the sun, all *Colias* drop to the ground; when the sun reappears, the butterflies warm up within a few minutes and are flying again. Thus habitat temperature, and particularly solar flux, seems a promising choice for an environmental factor importantly affecting the functioning of the enzyme $\alpha$-GPdH.

Temperature has proven a fortunate choice, as the biochemical behavior of enzyme alleles of $\alpha$-GPdH in *Colias* is indeed differently affected by reaction temperature (Johnson, 1976a).

Finally, the available habitats of the Colorado *Colias* encompass many very different thermal environments. The alpine species *C. meadii* is typically found on tundra, above timberline (~12,000 ft). The montane species *C. alexandra* occurs in montane open valleys at elevations of about 9,500 ft. *Colias scudderi* lives in montane habitats of 9,000–11,000 ft in conjunction with willow (the other species are restricted to legumes). The lowland species complex *C. philodice-C. eurytheme* occurs as an agricultural pest in lowland farmland from 5,000–8,000 ft. Transitional populations occasionally occur in which populations of the lowland complex occupy montane meadows, or in which the alpine species occurs in montane habitat.

## POLYMORPHISM AT THE $\alpha$-GPDH LOCUS IN *COLIAS*

Polymorphism at the $\alpha$-GPdH locus was examined in 18 populations over a period of five years (Johnson, 1976a). Two variant forms were detected by analysis of population samples on 7% polyacrylamide gels [there is reason to believe that additional "hidden" alleles exist which are not detected by this approach (Johnson, 1971, 1975, 1976b)]. Chemical characterization of the two allozymes indicates that the same homologous alleles occur in each of the five species examined. The two alleles segregate in crosses in a mendelian manner, and heterozygous individuals can be shown to possess three electrophoretic bands (the middle band being a heterodimer or hybrid molecule). The substrate binding kinetics of the two forms are significantly different. In particular, the faster-migrating variant enzyme binds substrate more effectively at 10°C (has a lower $S_{0.5}$) while the slower-migrating variant is more effective at 30°C. As this corresponds roughly to the thermal range of the butterfly habitat during the *Colias* flight season, this difference between the alleles is likely to be of adaptive significance.

When polymorphism for $\alpha$-GPdH is compared for the 18 populations, a significant pattern is evident (Table 1, after Johnson, 1976a): all nine montane populations are quite polymorphic (heterozygosity greater than 10%), while alpine or lowland populations are far less variable. This is true within species as well as between them. Thus alpine populations of *C. meadii* are not variable at this locus, while montane populations are. The result seems quite general over the five species: for 162 alpine individuals, heterozygosity for $\alpha$-GPdH averaged 6%; for 292 montane individuals, average heterozygosity was 35%; for 174 lowland agricultural individuals, average heterozygosity was 7%.

This pattern of genetic variability is consistent with what we know of the thermal nature of these habitats. While the montane valleys may get quite cold at night, they offer warmer habitats during the day than the wind-swept open tundra of the high alpine populations. It is in the colder high altitude populations that the fast allele predominates, and it is the fast allele which is the most effective binder of substrate at low temperature. Montane environments are

TABLE 1. Patterns of $\alpha$-glycerophosphate dehydrogenase polymorphism in several species of *Colias*.

| Species | Location (Altitude) | Habitat | Number of Individuals Analyzed | Heterozygosity of $\alpha$-GPdH |
|---|---|---|---|---|
| *C. meadii* | Cumberland Pass (12,000′) | Alpine | 48 | 0.04 |
| | Uncompahgre Peak (12,500′) | Alpine | 23 | 0.04 |
| | Mesa Seco (12,200′) | Alpine | 30 | 0.17 |
| | Upper Cement Creek (12,000′) | Alpine | 36 | 0.03 |
| | Upper Queen Basin (12,100′) | Alpine | 18 | 0 |
| | Copper Creek (10,200′) | Montane | 16 | 0.50 |
| | Los Piños Pass (10,500′) | Montane | 20 | 0.45 |
| *C. scudderi* | Upper Cement Creek (11,700′) | Alpine | 7 | 0 |
| | Lower Cement Creek (9,800′) | Montane | 21 | 0.24 |
| | Taylor Park (10,500′) | Montane | 32 | 0.17 |
| *C. alexandra* | East River (9,500′) | Montane | 76 | 0.42 |
| | Brush Creek (9,400′) | Montane | 40 | 0.45 |
| *C. philodice* | Slate River (9,100′) | Montane | 46 | 0.39 |
| | Lower Cement Creek (9,400′) | Montane | 22 | 0.27 |
| | Hotchkiss (5,000′) | Agricultural | 89 | 0.10 |
| | St. Louis, Mo. | Agricultural | 40 | 0.05 |
| *C. eurytheme* | Lower Cement Creek (9,200′) | Montane | 19 | 0.21 |
| | Los Baños, Calif. (500′) | Agricultural | 45 | 0.02 |

consistently less predictable in their thermal extremes than are alpine areas, so that the pattern as well as the range of the two habitats differ.

The general results are thus quite consistent with the hypothesis that enzyme polymorphism at the $\alpha$-GPdH locus reflects adaptation to a heterogeneous thermal environment. This hypothesis makes a clear and testable prediction: when single populations occupy diverse habitats, different portions of the population should experience very different selection. Thus, for example, a number of populations of *C. meadii* are known which occur right at timberline. Portions of these populations live in alpine habitats, while other portions extend down into montane habitats. The genetic structure of one such population at Mesa Seco has been studied intensively (Watt et al., 1976), and from their mark-release-recapture studies it appears that the population is genetically continuous, with at least a few individuals passing along its entire length each generation. Thus migration within the population would be expected to render it genetically uniform—unless very strong differential selection were acting upon the two alleles.

When $\alpha$-GPdH polymorphism is examined along a transect from alpine to montane areas within the Mesa Seco population, the transect is *not* uniform (Table 2). A pronounced *cline* in heterozygosity is seen: the alpine sites are essentially monomorphic for the fast allele, as observed previously, but the slow allele becomes increasingly more common as lower sites are examined. Only in the highly heterozygous lower sites are the two alleles in Hardy-Weinberg equilibrium; the higher sites show large deficiencies in the slow homozygote.

Again, this result seems quite general. When the same population was sam-

TABLE 2. $\alpha$-GPdH polymorphism along a transect through the Mesa Seco population.

| Year | Site Number | Altitude of Site (ft) | Number of Individuals Analyzed | Observed Frequency of $\alpha$-GPdH Heterozygotes |
|---|---|---|---|---|
| 1971 | 13 | 12,200 | 21 | 0.19 |
| | 13a | 12,000 | 9 | 0.11 |
| | 12 | 11,500 | 24 | 0.21 |
| | 10 | 11,000 | 21 | 0.43 |
| | 9 | 10,800 | 5 | 0.80 |
| 1973 | 13 | 12,200 | 40 | 0.11 |
| | 12 | 11,500 | 40 | 0.21 |
| | 11 | 11,300 | 40 | 0.38 |
| | 10 | 11,000 | 40 | 0.43 |
| | 9 | 10,800 | 20 | 0.47 |

pled again two years later, an identical cline was seen. Two other timberline populations, geographically quite distant, exhibit similar alpine-montane clines in $\alpha$-GPdH heterozygosity.

Thus variation at this locus is in all respects consistent with an adaptive hypothesis. One cannot, of course, rule out the possibility that selection actually is occurring on some other locus that we don't know about, and that $\alpha$-GPdH is simply linked to it. This argument is something like invoking divine intervention—it can be used to explain anything, and it is never possible to falsify it. However, if linkage forms the basis for the observed $\alpha$-GPdH polymorphism, it is remarkably fortunate that it has produced such a functionally suitable distribution of alleles!

## VARIATION AT OTHER LOCI—THE MULTI-LOCUS PROBLEM

While a single locus approach such as described above accounts reasonably well for the genetic polymorphism seen at the $\alpha$-GPdH locus, the result need not be general. Species of the genus *Colias* exhibit high levels of genetic variability at many enzyme loci (Table 3), and we have accounted for only one. What of the others?

To account for the generalized occurrence of enzyme polymorphism, one of two hypotheses is usually advanced. They are both fundamentally single-locus hypotheses. One is the hypothesis which we have used to account for the $\alpha$-GPdH variation: a heterogeneous environment selecting for different alleles under different circumstances. Similarly contrasting environmental influences are known to produce a polymorphism for sickle-cell hemoglobin in man, and have been implicated in lactate dehydrogenase (LdH) polymorphism in fish and alcohol dehydrogenase (AdH) polymorphism in *Drosophila*. However, if such single-locus explanations provide the basis for most of the polymorphic enzyme variation which is being reported, then we shall have to do a great deal of work to document this fact!

An alternative hypothesis is that of molecular overdominance: hybrid enzyme molecules (formed from subunits of both parental types) are viewed as in-

TABLE 3. Observed heterozygosity in eight *Colias* populations ($N > 40$).

| | *C. philodice* | | *C. scudderi* | *C. alexandra* | | *C. meadii* | | |
|---|---|---|---|---|---|---|---|---|
| Locus | Slate River (montane) | Hotchkiss (lowland) | Cement Creek (montane) | East River (montane) | Cement Creek (montane) | Cumberland (alpine) | Mesa Seco (alpine) | Mesa Seco |
| α-GPdH | 0.40 | 0.10 | 0 | 0.17 | 0.15 | 0.05 | 0.11 | 0.43 |
| G6PdH | 0 | 0.05 | 0.25 | 0.35 | 0.38 | 0.08 | 0.55 | 0.85 |
| MdH-1 | 0.10 | 0 | 0 | 0.05 | 0 | 0 | 0 | 0 |
| MdH-2 | 0.85 | 0.67 | 0 | 0.30 | 0.38 | 0.13 | 0.15 | 0.45 |
| ME | 0 | 0.20 | 0.25 | 0 | 0.08 | 0 | 0.05 | 0.15 |
| FUM | 0.11 | 0.45 | 0.20 | 0.05 | 0.08 | 0 | 0 | 0.55 |
| PGM | 0.45 | 0.35 | 0.10 | 0.11 | 0.18 | 0.28 | 0.65 | 0.65 |
| TPI | 0.35 | 0.15 | 0.25 | 0.20 | 0 | 0.15 | 0.05 | 0.05 |
| EST-1 | 0 | 0 | 0 | 0 | 0 | 0.05 | 0 | 0 |
| EST-2 | 0.65 | 0.50 | 0.15 | — | — | 0.20 | — | 0 |
| $\overline{HET}$ | 0.29 | 0.25 | 0.12 | 0.14 | 0.14 | 0.09 | 0.17 | 0.31 |

trinsically more stable or kinetically superior. This functional superiority produces a direct heterosis, and because heterozygotes are always at an advantage, high levels of polymorphism result. This hypothesis has great difficulty, however, in accounting for polymorphisms at loci of monomeric enzymes without multiple subunit structures.

Thus neither of these single-locus hypotheses is particularly satisfactory. I believe that the reason for this lies less with the hypotheses than with the question they address. The key is in realizing that α-GPdH, LdH, AdH and hemoglobin are atypical enzymes. Each involves a discrete physiological function directly affected by environmentally-imposed reaction conditions (Johnson, 1973b); it is not unreasonable that a single-locus hypothesis would be satisfactory in these cases. However, this is true of few of the other loci of Table 3. Most of the polymorphic enzymes are intimately involved in intermediary metabolism, and a change in the activity of one may influence the functioning of many others. Thus a change in hexokinase, which generates glucose-6-phosphate, cannot help but effect the reactions of phosphoglucomutase, phosphoglucoisomerase, and glucose-6-phosphate dehydrogenase, all of which use glucose-6-phosphate as a substrate. It seems likely that only a multi-locus hypothesis will be able to account for variation among such loci. In this regard it is worth noting that polymorphic variation occurs primarily at regulatory (rate limiting) steps in intermediary metabolism (Johnson, 1974). This finding is very widespread and quite general (the matter is extensively reviewed in Powell, 1975). This is a pattern which one would expect only if selection were acting on *the integrated metabolic phenotype* rather than on *individual loci per se*.

One multi-locus hypothesis which seems to me very attractive is that enzyme polymorphism is selected at regulatory loci so as to buffer these reactions from environmental perturbation (Johnson, 1976c). To maintain metabolic integration

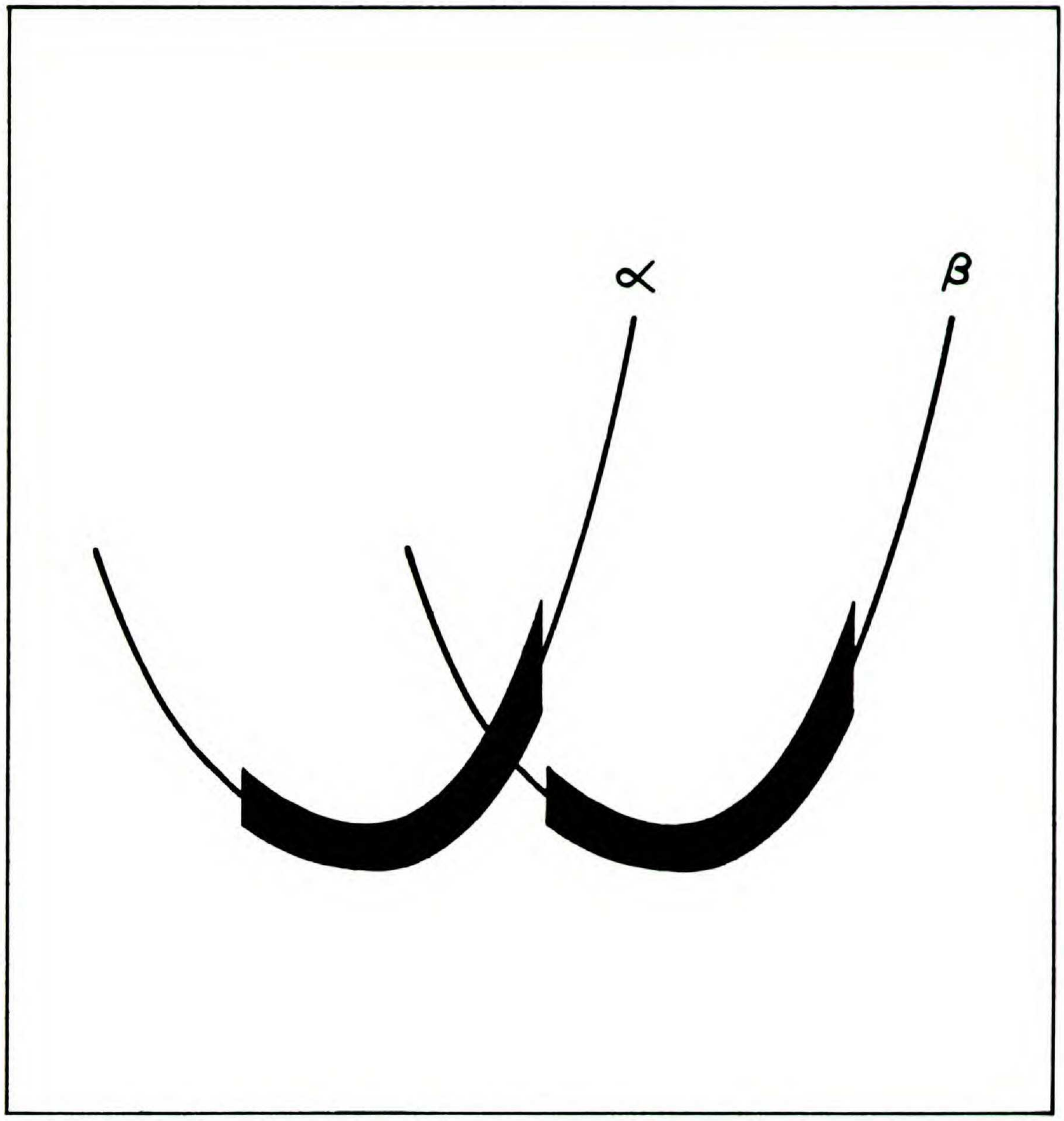


FIGURE 1. Strategies of metabolic regulation. At low temperatures form $\alpha$ has the lower $S_{0.5}$ ($K_m$) and binds most of the available substrate; both forms have allosteric sensitivity over the same range as their optimal affinity.

in a variable environment is of major evolutionary importance, as many metabolic control systems are interrelated. Yet critical reactions may respond quite differently to changes in temperature, etc. It may be of significant adaptive advantage to be able to maintain a constant relationship among critical regulatory reactions (Hochachka & Somero, 1973).

It is easy to envision molecular mechanisms which would produce such a homeostasis. Figure 1 provides one example. If the activities of two alleles respond differently to a habitat variable such as temperature (and such differences are well documented for tissue-specific isozymes), then the functional displacement with respect to temperature can buffer the coordination of metabolism from the effects of a temperature change. Any one allele of a regulatory enzyme

can exhibit a low $S_{0.5}$ (e.g. bind substrate well) only over a relatively narrow temperature range. This thermal sensitivity is an inevitable result of the requirement that regulatory enzymes be structurally flexible enough to be sensitive to allosteric "effectors" (low molecular weight molecules such as ATP whose binding acts as a metabolic signal). Over a broad temperature range a regulatory reaction cannot maintain a constant affinity for substrate and a constant binding affinity for effector molecules. As a result, it is difficult to maintain coordinate regulation with respect to other pathways catalyzed by different proteins responding differently to the change. A heterozygous individual, however, has *two* allelic forms present in each cell. In the example of Fig. 1, the $\alpha$ allele has the stronger binding affinity for substrate (lower $S_{0.5}$) at lower temperatures. When substrate concentrations are low, which is typically the case, only the $\alpha$ form will bind substrate at low temperature, and it will determine the reaction rate. Were it the only form present, the rate of binding would change at higher temperature. However, because of the functional displacement of the $\beta$ allele, the $\beta$ allele has the lower $S_{0.5}$ at higher temperatures. As a result, it is the $\beta$ form which binds the substrate at these temperatures—and the *realized* binding affinities have not changed over the broad range of temperatures!

This model suggests that heterozygotes are not overdominant so much as conditionally hemizygous, and that it is the very difference between the alleles which produces the adaptive advantage. Polymorphism is seen as a genetic strategy for maintaining metabolic integration in the face of environmental heterogeneity.

## Multiple Loci in *Colias*

The highly coordinated nature of intermediary metabolism suggests that if polymorphic alleles at regulatory enzyme loci are functionally different, then the particular allele present at one locus will importantly affect the activity of many other reactions. Thus if an individual possesses a low-temperature allele at one locus, rather than a higher-temperature form, then it makes a difference which functional forms are present at other regulatory loci. If a network of related regulatory enzymes are all optimally suited to low temperature, then one may speak of a *metabolic phenotype* adapted to these conditions. It is at this level that selection acts—on the expressed phenotype of individuals, rather than upon individual loci. Because the particular functional variant occurring at each regulatory locus influences the physiological state of the individual, *selection on metabolic phenotypes implies selection on allozymic genotypes.* To understand enzyme polymorphism, then, it will be necessary to simultaneously characterize a variety of loci in each sampled individual of a population.

Such a study is best carried out within a single population, to eliminate the possibility that differences in allele frequencies arise from demographic complications. If different loci sampled from the same individuals exhibit different patterns of allele frequency, then the result may not be attributed simply to migration or habitat selection by mobile adults.

For such a genotypic comparison, 14 loci of *C. meadii* were characterized, each individual butterfly being tested for all 14 loci. The Mesa Seco population was selected for the study, and samples were collected at each of five sites along

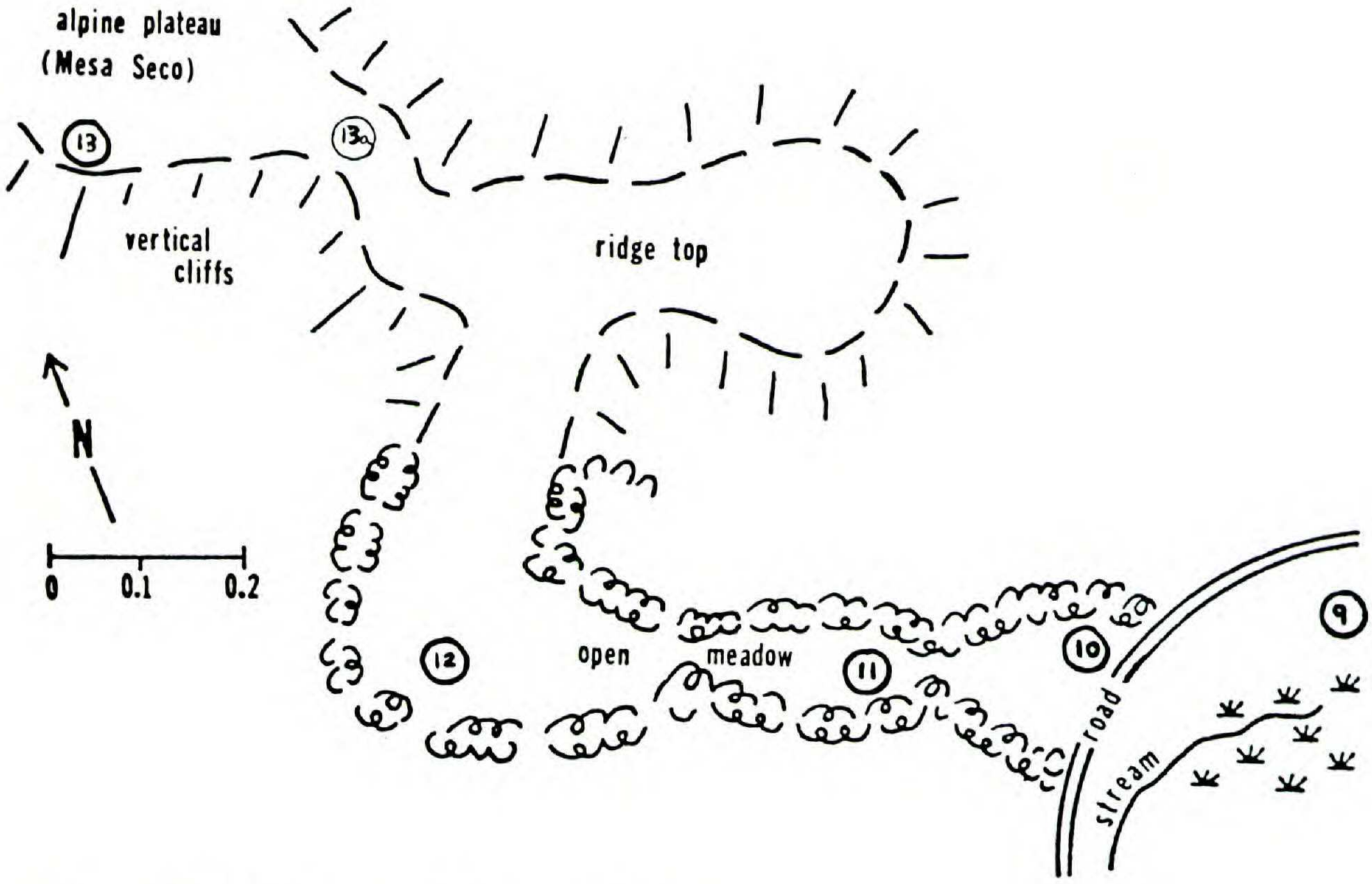


FIGURE 2. Mesa Seco population sample sites.

a transect from alpine to montane: an alpine site (#13), timberline (#12), montane forest (#11), montane forest-meadow boundary (#10), and open montane meadow (#9) (Fig. 2). The total distance traversed by the transect was about 2 miles, and the elevational difference about 2,000 ft. The population size was known from previous mark-release-recapture studies to exceed 1,000 individuals (one generation per year) and seemed to maintain approximately the same numbers from year to year. Individuals are seen to exchange between adjacent sites at a frequency of about 10%, although little or no exchange is seen between the more distal sites.

What then of the genotypes? A typical data set is presented in Table 4, that for the sample from site #9. A most startling relationship is apparent! Of 20 individuals, fully 9 appear to have highly organized genotypes: 5 individuals have identical alleles at each of 7 of the 12 loci, 4 others are identical at 6 loci. The odds that an individual would have such coordination are very low. For the first group of Table 4, the common genotype and associated allele frequencies are:

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Genotype | B | — | C | —/C | A/B | A | B | — | — | — | — | B |
| Allele frequency | 0.79 | — | 0.50 | 0.84 | 0.80 | 0.25 | 0.37 | — | — | — | — | 0.61 |

The joint probability that one individual will have this genotype is the product of the allele frequencies, $P = 0.02$. The probability that five individuals would possess this genotype by chance alone is only $(20/5)\ p^5q^{15}$, or $3.66 \times 10^{-5}$!

Two such commonly recurring genotypes are apparent in the sample from site #9, one involving six loci and one involving seven. The implication is very strong that they reflect selection for particular constellations of alleles.

TABLE 4. Genotypes at meadow site (#9).

| Individual Number | α-GPdH | EST-1 | EST-2 & 3 | AK-1 | AK-2 | G6PdH | HK-1 | HK-2 | ME | PGM | FUM | MdH-1 | MdH-2 | TPI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | B | — | C | B | C | A/B | A | — | B | A/C | B | — | A | B |
| 2 | B | C | C | A | C | A/B | A | — | B | B | C/D | C | A | B |
| 3 | B | A/B | C | B | C | A/B | A | — | — | B/C | C/E | C | A/B | B |
| 4 | B | — | C | A | C | A/B | A | — | B | A/B | B/D | B | A/B | B |
| 5 | B | B | C | A | C | A/B | A | — | B | B | A/B | B | A/B | B |
| 6 | B | C | D | A | C | A/B | A | B | D | B/D | C | C | B | B |
| 7 | B | A/C | D | A | C | A/B | A | B | D | A/C | C/D | A | B | A |
| 8 | B | A/B | D | A | C | A/B | A | B | D | B/C | D | C | B | A |
| 9 | B | C | D | A | C | A/B | A | B | D | B | D | — | B | C |
| 10 | B | B | C | B | C | A/B | A | B | B | B/D | D | C | B | B |
| 11 | B | A/C | D | A | C | B | A | C | C | A/C | D | C | B | B |
| 12 | C | A/C | A/D | B | — | B | C | B | E | B/D | A | — | A | C |
| 13 | B | — | B/D | A | C | B | A | B | C | A | C | — | B/C | B |
| 14 | B | A/C | C/D | — | — | A/B | C | B | E | B/C | D | C | B | C |
| 15 | B | A/B | — | A | C | A/B | A | B | B | A | C | — | B | — |
| 16 | — | — | C | B | C | A | A | B | C | A/C | B/C | A | B | — |
| 17 | A/B | A/B | C | A | C | A/B | A | B | B | B | D | C | B | A |
| 18 | A/B | — | B | A | C | A/B | A | B | C | C/D | B | B | B | A |
| 19 | A | — | C | B | — | A/B | A | B | C | B | D | C | B | B |
| 20 | B | C | C | B | — | A/B | A | B | D | A/D | D | C | B | B |

When the other sites are examined, similar results are obtained: highly organized genotypes involving more than half the examined loci repeatedly occur at high frequency. However, these genotypic combinations are different for each site! In the face of the observed adult migration, this is a remarkable result. None of the genotypic combinations common at one site are ever observed at any other. These results are summarized in Table 5. In this table, loci where more than one genotype occurs among the group are symbolized by a dash; invariant loci (identical for the entire sample) are indicated by parentheses.

The genotypic organization seen in the results of Table 5 clearly relates to the overall metabolic phenotype, as it involves almost exclusively regulatory as opposed to nonregulatory loci. Only those reactions which significantly affect the rate of intermediary metabolism seem to be included in the organized genotypes.

I have considered $\alpha$-GPdH and the esterases separately, as their functions are individually relatable to habitat factors such as temperature or secondary plant compounds. The genotypes of these loci also appear highly correlated with the localized habitat.

The unavoidable conclusion which one must draw from these results is that organized genotypes indeed exist in natural populations, apparently maintained by selection in the face of significant migration.

## POLYMORPHISM AS GENETIC STRATEGY

The genotypic associations described above suggest rather strong selection. For the genotypes of site #9, the indicated fitness (expected genotypic frequency/observed) is about 0.10. This seems very strong selection, and it raises the question of how the genetic and population structure of *C. meadii* has evolved to cope with what appear to be stringent environmental constraints.

The observed properties of the *C. meadii* genetic system are: (1) It involves a very large number of small chromosomes ($2n = 62$); (2) Although each female will lay several hundred eggs during a yearly flight season, population sizes remain relatively constant (this suggests a mean zygotic fitness of the order of 0.005); (3) Mark-release-recapture studies indicate low adult mortality, suggesting that selection is primarily at the larval stage; (4) Mating appears to be panmictic within local populations; (5) Members of individual subpopulations appear to be quite sedentary: while some individuals may forage for several hundred meters, the distribution of most adult individuals appears localized to portions of the cline; (6) Unlike alpine populations of *C. meadii* studied at other localities (where there is little exchange between subpopulations), there is significant exchange between adjacent subpopulations of the Mesa Seco population.

The observed genotypic associations may be maintained in such a genetic system by at least two very different genetic strategies. One strategy is that of linkage. If the key loci are tightly linked, then the observed high linkage disequilibrium would be an inevitable result of the low recombination fraction between them (Allard, 1975). Such a hypothesis implies that the subpopulations along the cline must be genetically isolated from one another, despite migration.

TABLE 5. Genotypic combinations specifying at least ½ of examined loci.

| | Environmental | | | Regulatory | | | | | | | Nonregulatory | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Site | αGPdH | EST-1 | EST-2 & 3 | AK-1 | AK-2 | G6PdH | HK-1 | HK-2 | ME | PGM | FUM | MdH-1 | MdH-2 | TPI | Freq. |
| 13 | A | — | A/— | B | ( ) | A/B | ( ) | B | B | — | — | — | B | A | 0.25 |
| | B | — | — | A | ( ) | B | ( ) | B | C | B | — | A | B | A | 0.33 |
| 12 | C | — | A/C | B | ( ) | A/B | ( ) | B | A | — | — | A | — | B | 0.20 |
| 11 | A | C | B | A | ( ) | A/B | ( ) | B | B | B/— | — | — | — | — | 0.20 |
| 10 | C | — | — | — | ( ) | ( ) | ( ) | B | C | A/— | — | — | — | — | 0.35 |
| | B | C | — | A | ( ) | ( ) | ( ) | B | B | — | — | — | — | B | 0.20 |
| | C | C | — | — | ( ) | ( ) | ( ) | B | D | — | — | — | — | — | 0.15 |
| 9 | B | — | C | — | ( ) | A/B | A | Null | B | — | — | — | A/— | B | 0.25 |
| | B | — | D | A | ( ) | A/B | A | B | D | — | — | — | B | — | 0.20 |

Otherwise genotypes common in one subpopulation would appear in adjacent ones. To maintain the observed genotypic discontinuities would require strong selection.

The alternative hypothesis is that the genetic system of *C. meadii* is analogous to that of the plants on which it feeds: that it utilizes the great segregational power of its high chromosome number to produce in each generation a wide array of genotypes. From this varied assortment a small fraction survives to become adults at any given site. Different sites might then select for different genotypes. Such a genetic strategy is highly flexible, being capable of reorganizing the genotypic constitution of a local subpopulation yearly. Such a strategy would constitute ideal adaptation to an unpredictably variable local habitat.

In this respect it is worth noting that efficient food processing is of paramount importance to *Colias* larvae, and that the regulatory loci of Table 5 encompass many of the key points in physiological regulation of intermediary metabolism.

Both segregation and linkage strategies imply a heterogeneous habitat and strong selection within the Mesa Seco population. It is possible to experimentally distinguish between them by reexamining these sites in subsequent years. The population subdivision suggested by a linkage strategy predicts temporal stability: the local genotypic combination should recur from year to year. In contrast to this, a segregational strategy implies temporal as well as spatial habitat variability: the local genotypic combinations may be quite different from year to year. The data are not yet available to distinguish between these two alternatives.

It is of interest to compare the results obtained for *C. meadii* with other *Colias* species. *Colias alexandra* is known from mark-release-recapture studies to be far more mobile (a single individual may be observed to move kilometers in a day). Preliminary data comparing samples collected at two sites six miles apart (from what appears to be a genetically continuous population) reveal a single common genotype (seven loci of fourteen, at a frequency of 20%) which is the same for both sites. This presumably reflects the fact that a broad montane valley is a more uniform habitat than an alpine-to-montane transect. The result is not unlike what one might have found looking only at the alpine population of *C. meadii*. Perhaps populations of *C. alexandra* occupying a more diverse array of habitats (if such could be identified) might exhibit more than one common genotype. Alternatively, other factors may limit its distribution to these open mountain valleys.

Several populations of *C. philodice* have been examined, both agricultural populations in alfalfa fields and montane meadow populations. In two respects these populations are quite different from the *C. meadii* and *C. alexandra* populations described above: (1) They exhibit a lesser number of alleles per locus; (2) I am unable to detect any common genotypic combinations involving a significant proportion of the examined loci. Unlike the indigenous species described above, *C. philodice* is very much a weedy species and is widespread in disturbed habitats. Perhaps it avoids the specialized habitat adaptations (and specialized genotypes) of the other species by a homeostatic biochemical strategy such as discussed earlier: A small number of functionally distinct alleles may act to

buffer key regulatory reactions so as to maintain constancy over a range of environmental variability. However, such a homeostatic strategy precludes fine-tuned adaptation to subtle environmental differences. The effect is to perpetuate a particular metabolic phenotype, while preserving its coherence. In the relatively similar habitats produced by human disturbance, such a "weedy" metabolic phenotype may be optimal. It permits a coherent metabolic phenotype with less selection than is required to maintain a fully organized genotype. The trade-off here is that such a "weedy" phenotype lacks genotypic flexibility, and will not readily alter in adaptation to particular circumstances. It is an approximate solution, arrived at cheaply.

It is clear that a great deal remains to be done to understand these patterns of genetic variation. To me, the most attractive conceptual framework within which to organize the findings discussed above is to view the patterns of enzyme polymorphism seen within *Colias* butterflies as adaptive strategies, which in each case *match the flexibility of the metabolic phenotype to the heterogeneity of the environment*. The detailed information needed to evaluate this interpretation involves both biochemical study of the differential functioning of allozymes, and far more extensive surveys of natural populations.

## Literature Cited


Allard, R. 1975. The mating system and microevolution. Genetics 79: 115–126.

Clarke, B. 1975. The contribution of ecological genetics to evolutionary theory: detecting the direct effects of natural selection on particular polymorphic loci. Genetics 79: 101–113.

Harris, H., D. Hopkinson & E. Robson. 1974. The incidence of rare alleles determining electrophoretic variants: Date on 43 enzyme loci in man. Ann. Hum. Genet. 37: 237–253.

Hochachka, P. & G. Somero. 1973. Strategies of Biochemical Adaptation. Saunders Co., Philadelphia, Pennsylvania.

Johnson, G. 1971. Analysis of enzyme variation in natural populations of the butterfly *Colias eurytheme*. Proc. Natl. Acad. U.S.A. 68: 997–1001.

———. 1973a. Enzyme polymorphism and biosystematics: The hypothesis of selective neutrality. Annual Rev. Ecol. Syst. 4: 93–116.

———. 1973b. The importance of substrate variability to enzyme polymorphism. Nature New Biol. 243: 151–153.

———. 1974. Enzyme polymorphism and metabolism. Science 184: 28–37.

———. 1975. Enzyme polymorphism and adaptation. Stadler Genet. Symp. 7: 91–116.

———. 1976a. Polymorphism and predictability at the α-glycerophosphate dehydrogenase locus in *Colias butterflies*: gradients in allele frequency within single populations. Biochem. Genet. 14: 403–424.

———. 1976b. Hidden alleles at the α-glycerophosphate dehydrogenase locus in *Colias* butterflies. Genetics 83: 149–167.

———. 1976c. Genetic polymorphism and enzyme function. *In* F. Ayala (editor), The Molecular Study of Biological Evolution. Sinauer Publ., Inc., Sunderland, Massachusetts.

Lewontin, R. 1974. The Genetic Basis of Evolutionary Change. Columbia University Press, New York.

Powell, J. 1975. Protein variation in natural populations of animals. Evol. Biol. 8: 79–119.

Selander, R. & W. Johnson. 1973. Genetic variation among vertebrate species. Annual Rev. Ecol. Syst. 4: 75–91.

Watt, W. 1968. Adaptive significance of pigment polymorphism in *Colias* butterflies. I. Variation of melanin pigment in relation to thermoregulation. Evolution 22: 437–458.

———, F. Chew, L. Snyder, A. Watt & D. Rothschild. 1976. Population structure of Pierid butterflies. I. Numbers and movements of some montane *Colias* species. [In preparation.]